Common Information for Public Health Electronic Reporting

CIPHER Guide Version 1.0
December 1999
http://www.cdc.gov/od/hissb/docs/cipher.htm

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

Introduction – Reference Guide

Version 1.0, December 1999

Intended Audience

This guide is intended for various audiences within the Centers for Disease Control and Prevention (CDC) as well as within the organizations of CDC's various partners. The intended audiences include public health surveillance program staff who are responsible for controlling and preventing the spread of diseases through ongoing surveillance activities, such as Epidemic Intelligence Service (EIS) officers, epidemiologists, and surveillance program officers/directors, as well as staff who are responsible for the design and development of software information systems to support surveillance activities.

This Guide is a work in progress, to which data elements will be added/adjusted as needed. CDC recognizes the importance of establishing a process to effectively manage and address issues noted by public health partners throughout the implementation of CIPHER standards. Thus, CDC is developing a process to receive comments from its public health partners, address the issues raised, and provide feedback on the disposition of the comments. In the interim, contact the HISSB Executive Secretariat with questions and comments regarding the standards detailed in this CIPHER Guide: hissb@cdc.gov, or phone 770-488-8301 or 770-488-8302.

Background

CDC and its partners in public health are responsible, through ongoing surveillance activities, for controlling and preventing the spread of diseases of importance to public health. Historically, those activities have been organized to address specific diseases or categories of disease. At the

federal level, this has been based on programmatic and budgetary authorities provided by Congress. These actions often influence state and local priority setting.

For many years, CDC has designed and implemented information systems to support surveillance for specific diseases and adverse health conditions, and made these systems available for state and local use. The systems have served the programmatic needs of disease surveillance well, and have provided the links between the CDC and its partners. Until now though, these systems have not been standardized. Each has its own user interface, its own scheme for coding and storing data, its own method for analyzing surveillance data, and its own method of reporting data to CDC. However, these systems are not used in isolation from one another. For example, an individual in a state health department may be responsible, on a daily basis, for entering and analyzing data on two or more of these systems. The lack of a standardized design increases both the time needed to train people to operate the systems and the likelihood of errors in entering or analyzing data. In addition, the lack of standardized coding methods makes the exchange of data between systems difficult and time-consuming.

In recent years, the CDC and its partners recognized the need to establish a coordinated approach to supporting surveillance activities. The CDC/ATSDR Steering Committee on Public Health Information and Surveillance System Development (also known as the Katz Committee) examined aspects of integrating CDC surveillance and information systems. They published a report in the spring of 1995, titled "Integrating Public Health Information and Surveillance Systems," that outlined three elements essential to the evolution of integrated health information systems: data standards; a communications infrastructure; and policy-level agreements on data access, sharing, and burden reduction. Based on the recommendations from this committee, the Health Information and Surveillance Systems Board (HISSB) was established in late 1995; its mission is to formulate and enact policy concerning the planning, development, maintenance, and use of integrated public health information and surveillance systems. HISSB sponsors activities to develop and/or evaluate mechanisms for integrating CDC surveillance information systems. One of the sponsored activities is the Integration Project. Its long-term goal is to develop a framework for complementary electronic information systems that automatically gather health data from a myriad of sources on a real-time basis, monitor the health of communities, facilitate the ongoing analysis of trends and detection of emerging public health problems, and provide information for setting public health policy.

HISSB Integration Project

The HISSB Integration Project is designed to bring consistency to various surveillance systems developed by CDC for state and local health department use. The goal of the project is to develop a framework—built around common data definitions, software architectures, user interfaces, and electronic data interchange (EDI) standards—that will allow the Divisions/Programs/Offices (DPOs) throughout CDC to develop similar information systems that more easily share data and software tools (or components) with one another.

Integration is specifically ***not*** a single software system being developed centrally for use by all programs within CDC. It is a set of specifications that will allow for more efficient software development and, as appropriate to the disease or condition, the linking of registries and more efficient sharing of information. DPOs will still have the flexibility and responsibility to develop their own solutions. They will also be encouraged to purchase and use commercial off-the-shelf software, or, when needed, to develop tools that can be shared by other surveillance systems. Beyond this, project participants representing the various DPOs will collectively determine what additional levels of integration are and are not appropriate.

The Integration Project is supported through many activities and the efforts of several workgroups that were formed to address various aspects of the project. The mission of the Common Information for Public Health Electronic Reporting (CIPHER) workgroup is to develop detailed standards for data elements required to support public health surveillance. Other committees are working on standards in other areas, such as system architectures, user interfaces, and electronic data interchange (EDI).

CIPHER Workgroup

The CIPHER workgroup was established to present, review, and evaluate various options for defining the common data elements. The workgroup's charge was to identify data elements commonly used throughout CDC programs, and to formulate a single definition and use for each element.

The CIPHER group recognized the need to have the identification include both data elements currently collected and used throughout CDC systems, and expansion to new areas such as behavioral risks. That is, a number of data elements and general data concepts were recognized as data whose standardization would benefit various CDC programs. To progress, the workgroup recognized the need to approach the task of formulating single definitions and uses for each of these elements in phases, and therefore categorized "the world" of needed data elements into groups, with various priorities attached to these groups.

The workgroup initially focused on data elements within the current core record of the National Electronic Telecommunications System for Surveillance (NETSS). NETSS is a computerized public health surveillance information system that provides the CDC with weekly data regarding cases of nationally notifiable diseases. Core surveillance data including date, county, age, sex, race/ethnicity, and some disease-specific epidemiological information, are transmitted electronically by the states and territories to CDC through NETSS.

A number of the data elements in the NETSS core deal with demographic data, and therefore are widely used throughout various CDC programs. Given their prevalence across various CDC systems, CIPHER initially focused on the NETSS core record variables.

Group meetings were held throughout the evaluation process. The workgroup reached a consensus on the definition of each data element. The group then developed the CIPHER Guide, which contains detailed information on each of these common data elements, including

the final CIPHER-approved data definition associated with each data element. The current version of this CIPHER Guide reflects the workgroup's initial focus on the NETSS core record. This Guide is an ongoing work in progress, and will be updated periodically as new data elements are added.

It will be important to move forward in defining the data elements which were categorized as next in priority, and which remain in queue. These data elements delve into broader and more challenging areas, such as clinical information, case management information, health care cost data, and behavioral and other risk factor data. A work plan is currently being developed by the HISSB Executive Secretariat to guide efforts in moving forward with the next categories of data elements, and ensuring that staff with pertinent skill sets in these broader areas are involved in the standardization process.

CIPHER Goals and Objectives

The goal of CIPHER is to establish standards for data used in surveillance, which will allow a consistent definition and a consistent implementation across programs. In addition, CIPHER's intent is to approach the implementation portion of a data element in various phases, including development of hardcopy data collection forms, development of electronic data entry screens, data storage/data associations, and electronic data interchange (EDI). (EDI is defined as the transfer of structured data by agreed message standards from computer to computer by electronic means. For more information, refer to the subsection on General Description of Electronic Data Interchange, in this Introduction.)

A common understanding of the definition, meaning, and implementation of the data elements used in supporting surveillance enhances communication across program areas and with CDC partners. This CIPHER Guide has been produced in an effort to provide this common understanding. The Guide includes a lexicon identifying the most important data elements to collect and use in surveillance, summaries of associated information such as vocabularies and valid values, and implementation sections detailing a majority of the implementation phases listed above. The availability of such comprehensive information facilitates the integration of information across program areas by creating comparable data.

CIPHER initially concentrated on the phases of implementation that involved hardcopy data collection forms, development of electronic data entry screens, and data storage/data associations. These phases of implementation were selected based on the Integration Project's initial focus on "tight integration" of CDC systems. That is, an initial charge of the Integration Project was to provide specifications which, once implemented by CDC programs, would achieve a common "look and feel" throughout CDC-supplied systems, from hardcopy form to electronic entry to data storage and associations. Such a common look and feel would answer the concerns that CDC has designed and implemented a number of disparate information systems to support surveillance for specific diseases and adverse health conditions for use by state and local health departments. State and local health departments have continued to ask CDC to adopt a coordinated, streamlined approach to surveillance information systems, and to reduce the CDC-imposed burden on surveillance operations. Thus, because tight integration of CDC systems was

an initial focus, CIPHER includes the standards for the hardcopy form, electronic entry screen, and data storage/association aspects of implementation.

The other equally important phase of implementation, electronic data interchange, focuses on aspects of "loose integration," in that EDI is concerned with exchanging messages in an open environment in which any organization (including state and local health departments that have developed their own integrated systems) can participate. The ability to share data electronically is indeed a future goal of the Integration Project. Hence, "loose integration" and EDI messaging will be a primary focus for CIPHER in the future, with a focus on the Health Level 7 (HL7) standard. It is important to note that while EDI messaging is designated as a future CIPHER task, aspects of EDI messaging have played, and will continue to play, a role in the CIPHER workgroup's charge to formulate a single definition and use for each element.

CIPHER is also defining standard algorithms for deriving the value of calculated data elements (e.g., the calculated age based on an event date and the person's date of birth). In addition, CIPHER is addressing global processing items such as audit data which provide a record of the historical changes made to system data.

To meet the goal of standardizing data elements, the following objectives have been defined:

Objective 1: Establish consistent definitions for information collected and used by surveillance systems throughout CDC.

Objective 2: Define standards for how questions are to be formatted and information is to be collected on surveillance case report forms.

Objective 3: Identify standards for the processing of data in electronic data entry systems, including value/label displays, reference table look-ups, and a minimum level of edit-checking.

Objective 4: Identify storage standards.

Objective 5: Provide guidance on Electronic Data Interchange (EDI).

Objective 6: Provide guidance on coding for the display of data in statistical analyses and reports.

CIPHER Process

The CIPHER process focused on defining individual data elements and concepts. Concepts can be either a group of data elements (e.g., all Date data elements) or a defined algorithm to provide standard processing of data elements.

The CIPHER workgroup met on a regular basis, initially focusing on the charge as described above, and on establishing a process to achieve its goals. The group identified the following as initial steps in meeting the CIPHER objectives. Note that a number of the first steps involved the study and research of various surveillance systems.

- Clearly illustrate the information necessary for surveillance, as represented by the various information systems
- Examine and understand the way information is defined, organized, and managed
- Determine data model hierarchy and metadata components used in modeling the participating programs' existing information systems, to better understand relationships between data elements within a system and across surveillance information systems
- Identify similarities and differences in information within and across CDC surveillance information systems
- Support the identification of the data elements that should be included in the common public health record
- Obtain a better understanding regarding surveillance information system data transfer protocols and electronic data interchange (EDI) improvement opportunities at the federal, state, and local levels.
- Prepare a draft outline of CIPHER report
- Create a sample CIPHER report for a data element
- Agree on CIPHER report format
- Outline the workgroup process and define phases
- Prioritize data elements/concepts and assign to one of three phases as described below:

Phase I Analyze the data elements within the current National Electronic Telecommunications System for Surveillance (NETSS) core record

Phase II Analyze the current NETSS extended records and other data elements that are shared among a number of (but perhaps not all) systems

Phase III Identify other CIPHER variables (e.g., clinical information, case management information, health care cost data, behavioral risk factors, etc.)

An important position taken by the CIPHER group up front was to adopt, whenever appropriate, standards from the CDC HISSB Standards and Liaison Committee as well as external standards development organizations (SDOs) such as Health Level 7 (HL7) or the American National Standards Institute (ANSI). If existing standards were not found or were not appropriate, it was determined that the CIPHER workgroup would work with the SDOs to include the additional elements and data element definitions needed for public health reporting.

Several members of CIPHER also participated in the Standards and Liaison Committee, thus enhancing communication between the two groups. There was also a mutual review process of each group's output documents. (Refer to Appendix III for more detailed information on the CIPHER review process.)

Reference Materials

Many resources were used in the preparation of the discussion documents to provide insight into the existing information systems, program needs, and industry standards. Workgroup participants provided information about their program needs and current systems, along with experiences from working with their partners. This information was supplemented with focused research, to provide a broad view of the topic.

Reference materials used included:

- Existing system documentation
- Existing surveillance forms
- Common Data Elements Implementation Guide from the HISSB Standards and Liaison Committee (http://www.cdc.gov/data/index.htm)
- Data models and metadata databases prepared for the following ten systems, which were among the first systems involved in the Integration Project:

  -National Electronic Telecommunication System for Surveillance (NETSS)
  -Surveillance component of the Sexually Transmitted Disease Management Information System (STD*MIS)
  -Surveillance component of the Tuberculosis Information Management System (TIMS)
  -HIV/AIDS Reporting System (HARS)
  -Adult Spectrum of HIV Disease (ASD)
  -Pediatric Spectrum of HIV Disease (PSD)
  -Supplement to HIV/AIDS Surveillance (SHAS)
  -Unexplained Deaths (UD) project
  -Active Bacterial Core (ABC) project
  -FoodNet

- Health Level 7 (HL7) version 2.3 and the version 3 Reference Information Model (RIM) (http://www.mcis.duke.edu/standards/HL7)
- American National Standards Institute (ANSI) X.12 standards (http://web.ansi.org/)
- Data Elements for Emergency Department Systems (DEEDS) Release 1.0 (http://www.cdc.gov/ncipc/pub-res/pdf/deeds.pdf)

- Health Insurance Portability and Accountability Act (HIPAA) web site (http://aspe.os.dhhs.gov/admnsimp/index.htm)

- Health Care Finance Administration (HCFA) web site (http://www.hcfa.gov)

CIPHER Guiding Principles and Global Definitions

Guiding Principles

Guiding principles are direction-setting guidelines for action. CIPHER established several during the workgroup sessions that steered the decisions on data element definitions. Consider these as principles that embody the foundation of the CIPHER definitions. The CIPHER guiding principles are as follows:

- **Establish a Single Definition**
  The most fundamental CIPHER guiding principle was to achieve standardization and to establish a single definition for each common data element. The definition had to address the full life cycle of the data element from data collection through database storage, and be a single solution.

- **Leverage Existing Standards**
  Reference existing standards, particularly those mandated by HIPAA and external standards development organizations (SDOs) such as HL7 or ANSI X12, as well as standards from the CDC HISSB Standards and Liaison Committee. Adopt existing standards whenever appropriate. If existing standards were not found or were not appropriate, work with the SDOs to include the additional elements and data element definitions needed for public health reporting.

- **Harmonize with the Public Health Surveillance Data Model**
  The Public Health Surveillance data model, under development as part of the Integration Project, is linked inextricably with all other components of the Integration Project, including CIPHER data standardization. Data models consist of definitions of the categories and types of data needed for a particular arena, with a diagram showing the relationships between them. Metadata databases contain information on the meaning of the variables, as well as how systems use the individual variables. Data models and metadata enable comparison of data variables across systems. The Public Health Surveillance data model forms the basis for building virtually all components of integrated surveillance systems. It describes, at the conceptual level, the purpose, data structure, and relationships within and between CDC systems, and provides a framework for public health data needs, ensuring consistent understanding and representation of data, and facilitating data sharing. The harmonization of data modeling and data standards efforts is an ongoing process.

- **Develop Standard Implementation**
  The public health surveillance process is complex and, until CIPHER, required the use of dozens of different forms, entry styles, and approaches to data transfer. CIPHER's intent is to go beyond supplying a single standard definition. All phases of implementation of a data element are to be defined and detailed, including development of Hardcopy Data Collection Forms, development of Electronic Data Entry Screens, Data Storage/Data Associations, and Electronic Data Interchange for data transfer.

- **Support Electronic Exchange**
  Surveillance software must easily exchange information electronically through standardized import and export features. Public health recognizes that the use of electronic data interchange (EDI) messages based on the Health Level 7 (HL7) standard is consistent with health industry practices.

- **Make Optimal Use of Resources**
  To operate effectively, information system software requires many supporting components, such as hardware, reference files, and support personnel. The design of the system affects the requirements for these resources and, when optimal, reduces the potential burden. CIPHER promoted a centralized management of system reference files and tables to ensure timely upkeep and to reduce costs (e.g., for license fees, storage space). Identification of common algorithms and requirements also supports the development of shared software components.

- **Promote Individual Case Reporting**
  CIPHER definitions support both individual and summary case reporting, as required by current practice and legislation. A goal of the workgroup was to encourage movement toward the sole use of individual case reports through implementation of the CIPHER standards.

Global Definitions

The CIPHER discussions on the initial data elements revealed several aspects common to each data definition. These items were identified as global definitions that apply to all data elements or to a group of data elements with common features.

Global Data Element Characteristics

This section summarizes the global definitions that focus on coding/storage. All CIPHER data elements are defined as character type fields. The use of character format applies to all data including categorical data elements with alphabetic codes as valid values, ordinal data elements with numeric codes as valid values, free-form entry data elements, and scalar data elements in which data are stored as numbers and refer to a measurement (such as age).

The exclusive use of character data allows for a consistent implementation of systems across a variety of platforms. Rather than relying on intrinsic data types, which might or might not be available in specific database management systems, virtually any sort of data that a surveillance system might wish to manage can reasonably be handled as character strings.

Some of the specific advantages to choosing character representation include the following:

- CIPHER standards are virtually guaranteed to be 'implementable' on all database management systems. (Exceptions would include graphical and multimedia data types.)

- Utilities that read through standard interfaces such as Open DataBase Connectivity (ODBC), as required in the National Electronic Disease Surveillance System (NEDSS) "Surveillance System Component Guide", will recognize how to treat the data they are receiving.

- By avoiding database specific types, such as date fields and numeric values, more flexible solutions such as dates with missing subcomponents, as well as automatically formatted phone numbers, may be considered and consistently implemented.

The potential drawbacks to using only character fields are few. For example, some additional data manipulation may have to be performed during analysis when executing numeric operations against scalar data elements that are stored in character format. However, CIPHER itself has defined only three scalar variables ('Stated Age', 'Calculated Age', and 'Number in Household'). All others are either categorical or ordinal. Further, when numeric calculations are required, most modern database management systems will perform either implicit on-the-fly conversions of data types to numeric, or will allow these conversions to be performed explicitly. For example, in the Statistical Analysis System (SAS), character strings will automatically be converted to numeric when used in statistical functions or arithmetic operations. In addition, the American National Standards Institute (ANSI) Structured Query Language,Version 2 (SQL2) provides this capability through the CAST, TO_INTEGER, or TO_FLOAT functions. Thus the potential analysis-oriented drawbacks to using only character fields are minimal and are out-weighed by the advantages cited above.

A summary of the global definitions that focus on coding/storage follows:

All Data Elements:

- The meaning and intent of each data element are defined
- All data elements are defined as character type fields. A blank value (all spaces) is the default.
- A missing value is represented as a blank
- Only raw data are stored in the database
- Only a valid value code associated with the data element can be entered and stored in the database
- Derived fields will be calculated as needed and not stored in the database
- Text descriptions of valid value codes are solely for cosmetic use and are not stored in the database
- Text descriptions of valid value codes will be obtained from system tables/files

Additional Global Definitions for Categorical Data Elements:

- Valid values will be alphabetic codes
- Yes/No indicators will all use the valid values of “Y” for yes and “N” for no

Additional Global Definitions for Ordinal Data Elements:

- Valid values will be numeric codes, stored in character format

Additional Global Definitions for Alphabetic and Alphanumeric Data Elements:

- Fields will left justify and pad with blanks
- Case sensitivity is retained from data entry through storage. Thus, the free-form entered data are stored EXACTLY as they are entered (UPPERCASE, lowercase, or a mixture of both).

Additional Global Definitions for Scalar Data Elements:

- Fields will right justify and fill with blanks with the exception of zip code-related fields
- There will be no leading zeroes for numeric values
- A value of all 9’s is not permitted (i.e., a value of 9999 for a four-digit element is invalid)

Global Format Characteristics for Categorical and Ordinal Data Elements

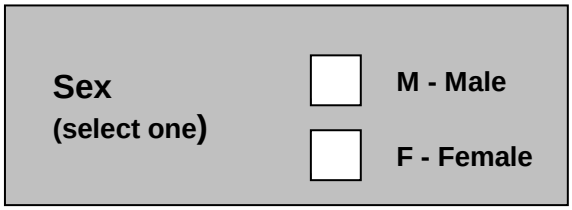
This section summarizes those guiding principles that focus on the appearance of all CIPHER data elements on hardcopy report forms, electronic data entry screens, and output reports.

Hardcopy Report Form:

- CIPHER data elements that are defined as coded variables, and therefore have a pre-defined list of valid value codes, are noted on the hardcopy form in a "check-box" format as described in the next two bulleted items. CIPHER data elements that are open "free-form" text type fields are to be noted on the hardcopy form by a blank line.

- The valid value codes and code descriptions associated with a CIPHER data element are included on the hardcopy report in check-box format unless the list is significantly large. (In such a case, the list of valid codes will be accessible through a separate hard reference.)

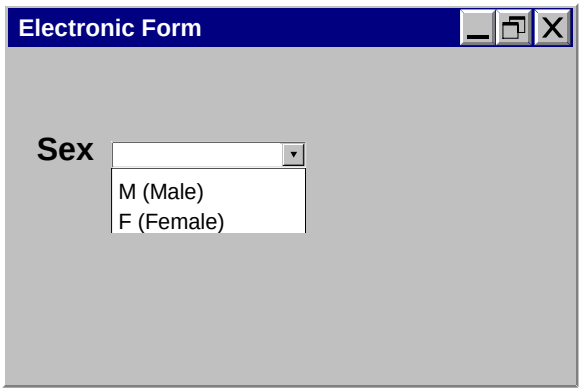
- The check-box format of the codes and descriptions will be formatted on the hardcopy report form as follows: "code value – label", (e.g., **F – Female)**. Consider the following example of a hardcopy report form containing a field for Sex:



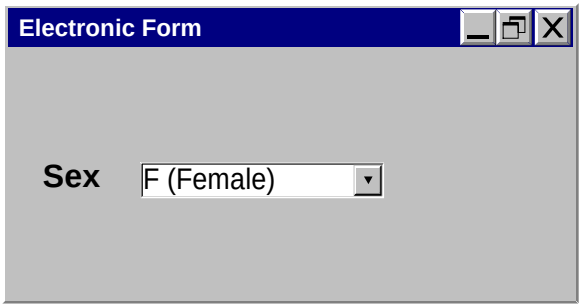
Electronic Form:

- CIPHER data elements that are defined as coded variables, and therefore have a pre-defined list of valid value codes and categories, will be displayed on the electronic entry screen along with a drop-down pick list. The pick list will be formatted in one of two ways as described below:

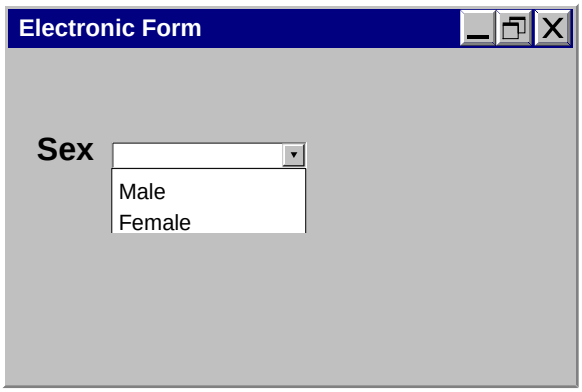
  - ***Code**-label* pick list format: Pick list displaying the valid codes, as well as the descriptive labels associated with the codes. The descriptive labels are displayed in parentheses. The code displayed to the left of the descriptive text reflects the data that are stored. Consider the following example of an entry screen for Sex in which the ***Code**-label* pick list format is applied:



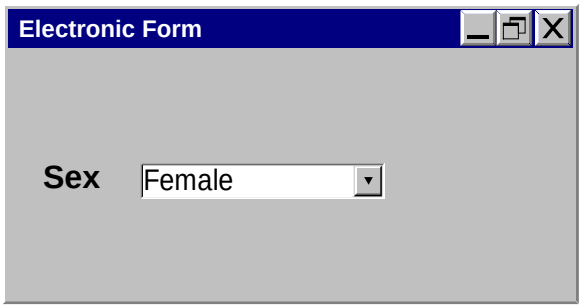
Once entered, the raw code will be displayed within the electronic field. In addition, the descriptive label will be displayed next to the stored data in parentheses. Consider the following example, in which F reflects the raw code stored within the Sex variable, and (Female) reflects the descriptive label.



- *Label only* pick list format: Pick list displaying only the descriptive labels. The associated codes that are stored in the database are not displayed or referenced. Consider the following example of an entry screen for Sex in which the *Label only* pick list format is applied:



Once a category is selected and entered, only the descriptive label will be displayed in the field. Consider the following example:



- The default pick list format, described above, is established by the user through a systems option feature (to be developed). The user is given the option to access either the ***Code**-label* format or the *Label only* format. The selected format is set as the default pick list format within that particular user's profile, and applies to all variables within all programs accessed by that user. The user may change the default pick list format at any time through a systems option.

  For the purposes of this document, the ***Code**-label* format is used in all electronic pick list illustrations.

- The list of valid values displayed in the pick list can be narrowed by having the user type in the first few characters of the entry text. This feature assists the entry operator with locating the desired entry value. For example, the user may enter the text **'303'** in a zip code field, and then access the pick list. The list of values noted in the pick list will contain all zip codes that start with the text **'303'**. This feature will be supported through a component (to be developed).

- To facilitate the data entry process, the entry field can be populated (automatically filled) with the valid value that most closely matches the text entered by the user. For example, if the entry operator key enters an **M** in the Sex field, the field will automatically be populated with **M (Male)**. This feature will be supported through a component (to be developed).

- CIPHER data elements that are open "free-form" text type fields will be displayed on the electronic entry screen along with a direct entry-type field.

Output Reports:

- CIPHER data elements included in electronic reports and/or hardcopy reports will be output in the ***Code****-label* default format. Consider the use of the ***Code****-label* in the following example of a one way frequency of Sex:

Sex	**Count**
M (Male)	56
F (Female)	44

General Description of Electronic Data Interchange

Electronic Data Interchange (EDI) is defined as the transfer of structured data by agreed message standards from computer to computer by electronic means. EDI is concerned with exchanging messages in an open environment in which any organization can participate. To operate in an open environment, no single organization can impose its own data definitions and standards on others. For EDI to work, the industry must agree on standard formats both for messages and for the way in which they are applied. Trading partners need not have identical information systems, because standardization makes the process independent of any specific hardware or software. Heterogeneous systems are linked through the use of a common format for data transfer. Using a standard to exchange data between systems saves time and money by eliminating the need to re-key data into multiple systems and/or to develop custom interfaces to enable two systems to exchange data.

EDI transactions can be divided into two components: message generation and communication. Translation software is usually required to interpret the data extracted from local systems. Commercial EDI software provides the mapping and translation functions for generating EDI messages. Mapping functions convert data from a local representation used by an EDI application to an internal representation of a standard EDI message, and vice versa. Translation functions convert an internal representation of a standard EDI form to an encoded standard EDI form and vice versa. Once an EDI message is generated and encoded, it must be communicated to a partner.

EDI standards specify data formats, but are independent of communications protocols. The delivery of standardized data to their destination may be achieved in a number of ways. The most common methods are: point-to-point or direct EDI through a direct link between two computers; value added networks (VANs) that provide an electronic mailbox with a multitude of services; and the Internet, through electronic mail or file transfer.

Many EDI standards exist. The two most widely used are the American National Standards Institute (ANSI) X12 and the EDIFACT (EDI For Administration, Commerce, and Transport) family of standards. CIPHER focused on the two health-specific EDI standards–X12.N (under the ANSI Insurance and Healthcare umbrella) and Health Level 7 (HL7)–because they are the most relevant to public health. Also, X12 and HL7 are likely to be adopted for electronic administrative health data as directed by the Health Insurance Portability and Accountability Act of 1996 (HIPAA).

HL7 is one of several ANSI-accredited Standards Development Organizations (SDOs) operating in the health care arena. Most SDOs produce standards (sometimes called specifications or protocols) for a particular domain (e.g., pharmacy, medical devices, imaging, or claims processing transactions). HL7's domain is clinical and administrative data. ANSI Accredited Standards Committee (ASC) X12.N standards are focused on the administrative and financial transactions needed to support the health care industry.

CIPHER Guide Overview

The CIPHER Guide contains the results of workgroup discussions and the document review process. It has a structured format to provide the reader with a consistent presentation of the topics covered for each data element.

Section Structure

Each section contains two major subsections: Summary and Implementation. They are complementary subsections and do not contain redundant information. There is a reference guide at the beginning of each section to provide the location of the page number of the major content headings.

This structure was designed to aid the various audiences in locating information of interest. These audiences include public health surveillance program staff who are responsible for controlling and preventing the spread of diseases through ongoing surveillance activities, such as Epidemic Intelligence Service (EIS) officers, epidemiologists, and surveillance program officers/directors, as well as program representatives who are responsible for the design and development of software information systems to support surveillance for specific diseases and adverse health conditions.

Summary Subsection Overview

The first subsection is the summary of the CIPHER definition for the data element or concept. It contains key components of the definition and provides a quick reference for all users. The Summary subsection contains the following components:

- Definition
- Data Storage and Field Values
- Missing Values
- Processing Overview
- EDI Summary
  (Note: As discussed in the section on CIPHER Goals and Objectives, EDI messaging is a future focus for CIPHER. Therefore, these subsections remain under construction as of the writing of this first version of the CIPHER Guide.)
- Discussion

Implementation Subsection Overview

The implementation of a data element definition can be described in various phases from design of the hardcopy data collection form to database storage.

The Implementation subsection contains the following components:

- Data Collection: Hardcopy Report Form
- Data Entry: Electronic Forms

- Data Processing: Validations and Edit Checks
- Data Processing: From Hardcopy to Storage
- Data Transmission: Electronic Data Interchange.
  (Note: As discussed in the section on CIPHER Goals and Objectives, EDI messaging is a future focus for CIPHER. Therefore, these subsections remain under construction as of the writing of this first version of the CIPHER Guide.)

Examples throughout the CIPHER Guide

It is important to acknowledge that a data concept may relate to more than one use of a data element. For example, "Address" is a data concept covered in this CIPHER Guide. There can be many uses of Address, such as address of subject of report, address of contact, and address of physician. Each use of address is associated with a specific data element, defined as an Address type field.

As described throughout this Introduction, the intent of the CIPHER Guide is to provide detailed information on the use of a data concept, and with that, the implementation of a specific data element associated with the data concept. Implementation covers development of hard copy data collection forms, development of electronic data entry screens, data storage/data associations, and electronic data interchange.

For the purposes of describing and illustrating the implementation of a data concept, the focus of the first iteration of this Guide is on data elements associated with the *subject of the report*. The implementation for other data elements that fall under the same data concept parallels the implementation described for the data element associated with the subject of the report. Thus, for example, the Address section of this CIPHER guide contains definitions on the general Address data concept, as well as specific information on the implementation of the data element defining the *Address of the subject of the report*. The implementation of other Address-related data elements can be patterned after this implementation.

In addition, CDC is developing a process to make available the definitions and variable names for the other uses of individual data concepts.

The electronic screens illustrated in figures throughout the CIPHER Guide were drafted by the CIPHER workgroup prior to the release of the CDC User Interface Style Guide. Refer to the draft CDC User Interface Style Guide for standards and guidelines on the development of CDC Windows and Web-based surveillance applications. That guide will provide further information and specifications on how these screens should appear to the user, and how the user will interact with the application.